DEPARTMENT OF STATISTICS

The University of Wisconsin
Madison, Wisconsin

TECHNICAL REPORT NO. 171

August 1968

PERCENTAGE POINTS OF THE NORMAL SCORE LAYER RANK TESTS FOR INDEPENDENCE AND EMPIRICAL POWER COMPARISONS.[1]

by

G.K. Bhattacharyya, Richard A. Johnson

and

H.R. Neave

1. This research was partially supported by the Office of Naval Research through contract Nonr 1202 (17) Project No. 042-222 and by the Wisconsin Alumni Research Foundation through Grants 67-08500-1 and 67-45435-1.

Summary

A class of nonparametric tests based on the third quadrant layer ranks has recently been studied by Woodworth ~~[13]~~ in connection with the problem of testing for independence in a bivariate distribution. In the present work, exact one-sided rejection regions are tabulated for the normal score layer rank test which is asymptotically locally most powerful for positive dependence in the bivariate normal distribution. The cut-off points are tabulated for sample sizes $n=4(1)9$ and significance levels $\alpha=.10$, .05, .025 and .01. Normal and Edgeworth approximations for the significance probabilities are also given. A simplified version of the normal score test is proposed and its rejection regions are tabulated. These tests are compared with the correlation coefficient test, Kendall's t test and Spearman's rank correlation test for independence by means of Monte Carlo evaluation of power employing 10,000 trials from each of three different types of bivariate distributions. Also included is a brief description of the computing aspects of the problem that may prove useful in similar studies. ( )

1. Introduction

Let $\underset{\sim}{Z}_i=(X_i,Y_i)$, $i=1,2,\ldots,n$ be a random sample from a bivariate distribution having the continuous cumulative distribution function (cdf) $F(x,y)$. The problem of testing for independence between X and Y, that is, of testing the null hypothesis $H_o:F(x,y)=G(x)H(y)$ against the alternatives of positive dependence has been studied extensively in the literature. The UMP unbiased test under the bivariate normal model is based on the sample correlation coefficient r. Various nonparametric tests have also been proposed. The two classical ones are the tests based on Spearman's rank correlation r_s and Kendall's t. A unified treatment of the asymptotic distribution theory of these and other rank order tests has been given by Hájek and Šidák [6]. Lehmann [10] gave some mathematical characterizations of positive dependence and demonstrated that r_s and Kendall's t possess desirable properties for specific types of positive dependence.

A new class of distribution-free tests of independence based upon the "3rd quadrant layer ranks" has been studied by Woodworth [13]. The 3rd quadrant layer rank of $\underset{\sim}{Z}_i$ is defined as the number of points (X_j-X_i,Y_j-Y_i), $1\leq j\leq n$, that lie in the closed 3rd quadrant. Let $X_{(1)}<X_{(2)}<\ldots<X_{(n)}$ and set $\underset{\sim}{Z}_{(i)}=(X_{(i)},Y_{[i]})$, $i=1,2,\ldots,n$ where $Y_{[i]}$ is the Y-component of the vector having the jth smallest X-component. The 3rd quadrant layer rank (henceforth to be called simply "layer rank")

$\ell(j)$ of $Z_{(j)}$ is then the rank of $Y_{[j]}$ among $Y_{[1]}, Y_{[2]}, \ldots, Y_{[j]}$.

The class of layer rank statistics considered in [13] has the structure $T_n = \sum_{j=1}^{n} E_n(\ell(j), j)$, where $E_n(i,j)$, $1 \le i \le j \le n$ is a triple sequence of constants representing the weight function associated with the layer ranks. It includes Kendall's t as a special case as can readily be seen by choosing $E_n(i,j)=i$. Another important member is the normal score layer rank statistic

$$T_n^{(1)} = \sum_{j=1}^{n} b_{nj} c_{j,\ell(j)} \tag{1.1}$$

where $c_{jk}=E(V_j^{(k)})$ with $V_j^{(1)}<\ldots<V_j^{(j)}$ an ordered sample of size j from the standard normal distribution and

$$b_{nj} = \begin{cases} 0 & \text{for } j=1 \\ c_{nj}-(j-1)^{-1}\sum_{i=1}^{j-1} c_{ni} & \text{for } 2 \le j \le n. \end{cases} \tag{1.2}$$

It is shown in [13] that for the alternatives of positive dependence in the bivariate normal family, the test which rejects H_o for large values of $T_n^{(1)}$ is asymptotically locally most powerful among all linear layer rank tests.

Although the asymptotic properties like the Pitman and the Bahadur efficiencies of the layer rank tests were studied in [13], no table of significance probabilities was given to aid in carrying out the tests and the only existing tables are for Kendall's t statistic [7,8]. In Section 2, we provide a table of upper 10%, 5%, 2.5% and 1% points for $T_n^{(1)}$ for sample

sizes n=4(1)9 as well as the normal and Edgeworth approximations to the significance probabilities. The handling of the distribution of $T_n^{(1)}$ is difficult due to the complicated form of sequence $\{b_{nj}\}$. Analogously to the Kendall's t statistic, a simplified version $T_n^{(2)}$ of the normal score layer rank test is constructed from $T_n^{(1)}$ by taking the b_{nj} identically equal to 1. Thus,

$$T_n^{(2)} = \sum_{j=1}^{n} c_{j,\ell(j)} \quad (1.3)$$

A table of upper percentage points of $T_n^{(2)}$ is also provided in Section 2.

The asymptotic relative efficiencies of the tests r, r_s, t and $T_n^{(1)}$ have been studied by Konijn [9], Bhuchongkul [2], Woodworth [13] and others. No information is, however, available about their performance in small and moderate sample sizes. Computation of the exact power being extremely difficult, we present in Section 3 tables of the empirical power of all these tests and $T_n^{(2)}$ under three bivariate distributions. From the enormous number of samples (10,000) used in the study, one would expect that the empirical values are fairly close to the exact power for the alternatives considered.

2. Percentage points of $T_n^{(1)}$ and $T_n^{(2)}$

Let $\underset{\sim}{\ell}=(\ell(1),\ell(2),\ldots,\ell(n))$ denote the vector of layer ranks of $(\underset{\sim}{Z}_{(1)},\underset{\sim}{Z}_{(2)},\ldots,\underset{\sim}{Z}_{(n)})$ and $\mathcal{C}(\underset{\sim}{\ell})$ the set of all possible layer rank vectors. Under H_0, the different components of $\underset{\sim}{\ell}$ are mutually independent and $\ell(j)$ is uniformly distributed over the set of integers $\{1,2,\ldots,j\}$, $1\leq j\leq n$ (cf. Lemma 1.1 of Brandoff-Nielsen and Sobel [1]). Thus the $n!$ vectors involved in the set $\mathcal{C}(\underset{\sim}{\ell})$ are equally likely under H_0. Determination of the significance points of a layer rank statistic T_n requires generation of the set $\mathcal{C}(\underset{\sim}{\ell})$, calculation of the values of T_n over the set and then the ordering of these values.

Certain techniques have made it possible to substantially reduce the computing time. Both $T_n^{(1)}$ and $T_n^{(2)}$ are handled simulteneously, thus avoiding regeneration of $\mathcal{C}(\underset{\sim}{\ell})$. Secondly, it is found that an auxiliary statistic $U_n(\underset{\sim}{\ell})=\sum_{j=1}^{n}\ell(j)$ (this is equivalent to Kendall's t) helps to generate the layer rank vectors $\underset{\sim}{\ell}$ leading to the high values of $T_n^{(1)}$ and $T_n^{(2)}$ due to the monotonic nature of the weight functions involved. This is made precise in the following lemma.

Lemma 2.1. For integers r, $n<r\leq n(n+1)/2$, define the subsets $S(r)$ of $\mathcal{C}(\underset{\sim}{\ell})$ by

$$S(r) = \{\underset{\sim}{\ell} : U_n(\underset{\sim}{\ell}) = r\}. \tag{2.1}$$

Then for $i=1,2$

$$(2.2) \qquad \max_{\underset{\sim}{\ell}\in S(r)} T_n^{(i)}(\underset{\sim}{\ell}) > \max_{\underset{\sim}{\ell}\in S(r-1)} T_n^{(i)}(\underset{\sim}{\ell}).$$

Proof. Consider, in particular, $T_n^{(1)}(\underset{\sim}{\ell}) = \sum_{j=1}^{n} b_{nj}c_{j,\ell(j)}$. For every $\underset{\sim}{\ell}_o \varepsilon S(r-1)$, there exists a layer rank vector $\underset{\sim}{\ell}_o' \varepsilon S(r)$ such that for some integer i $(1 \le i \le n)$, $\ell_o(i)+1=\ell_o'(i)$ and $\ell_o(j)=\ell_o'(j)$ for all $j=1,2,\ldots,n$ with $j \ne 1$. The sequence $\{b_{nj}\}$ is non-negative and non-decreasing in j. Also, for each fixed j, the sequence $\{c_{ji}\}$ is increasing in i. This entails $T_n^{(1)}(\underset{\sim}{\ell}_o) < T_n^{(1)}(\underset{\sim}{\ell}_o')$ and hence (2.2) follows for i=1. The same argument applies for $T_n^{(2)}$ and this completes the proof.

To avoid generating all the n! elements of $\mathcal{C}(\underset{\sim}{\ell})$, the layer rank vectors are generated in decreasing order of $U_n(\underset{\sim}{\ell})$. For instance with n=5, we start from $S(15)=\{(1,2,3,4,5)\}$ and then generate the set $S(14)=\{(1,2,3,4,4), (1,2,3,3,5), (1,2,2,4,5), (1,1,3,4,5)\}$ followed by the set S(13) and so on. For each n and each test statistic $T_n^{(i)}$, a reference number C_i is chosen which is sure to lie below the 10% cut off point (hence also below the 5%, 2.5% and 1% points). The sets S(r) are generated in decreasing order of r and for each $\max_{\underset{\sim}{\ell}\varepsilon S(r)} T_n^{(i)}(\underset{\sim}{\ell})$ is computed. As soon as this becomes less than $\min(C_1,C_2)$, the program is terminated. (2.2) ensures that the set of all the layer rank vectors generated under this scheme contain the critical regions of the desired sizes.

Even with the above technique, more than n!/10 values of $T_n^{(i)}$ are calculated and these are to be ordered. The best ordering program requires computing time proportional to NlogN,

where N numbers are being ordered. An even more serious problem is the lack of sufficient storage space. With n=9, for instance, about 45,000 values are to be ordered for each statistic. It is frequently experienced that the computation of the test statistics takes much less computer time than the printing out of the computed values. The program is therefore run in two stages. In the first stage, the computed values of each statistic are grouped into a histogram having 1,000 cells. The intervals containing the four percentage points under consideration are located. In the final run, the values within these four intervals are ordered and printed out and from these, the exact cut off points are determined. About 15 minutes were required for the case n=9.

The 10%, 5%, 2.5% and 1% upper cut off points of $T_n^{(1)}$ and $T_n^{(2)}$ for n=4(1)9 are presented in Table 1. In some cases the nominal α can only be attained through randomization on the boundary. An entry within braces represents the nonrandomized significance probability (in percentage) corresponding to the cut off point marked by an asterisk immediately above it. The number immediately below is the next lower value of the test statistic which is to be randomized. For instance, with n=4, α=.10, the test based on $T_n^{(1)}$ is the randomized test: reject H_o with probability 1 if $T_n^{(1)} \geq 1.8260$ and reject H_o with probability (.10-1/12) if $T_n^{(1)}=1.8122$.

Table 1 The upper percentage points of the null distributions of the normal score layer rank statistics

	$T_n^{(1)}$				$T_n^{(2)}$			
n \ α	.10	.05	.025	.01	.10	.05	.025	.01
4	*1.8260 ($8\frac{1}{3}$%) 1.8122	*2.6386 ($4\frac{1}{6}$%) 1.8260	—	—	*1.7075 ($8\frac{1}{3}$%) 1.5936	*2.4398 ($4\frac{1}{6}$%) 1.7075	—	—
5	2.1569	2.4578	3.0938	*3.8475 (5/6%) 3.1459	1.9102	2.4398	2.8704	*3.6028 (5/6%) 2.9349
6	2.3357	2.9985	3.5295	*3.9496 (35/36%) 3.8756	2.1464	2.6973	3.1775	*3.7416 (35/36%) 3.7071
7	2.6001	3.3064	3.9001	*4.6003 (125/125%) 4.5954	2.3908	3.0500	3.6016	*4.2271 (125/126%) 4.2130
8	2.8968	3.6579	4.2890	*5.0535 (2015/2016%) 5.0504	2.6434	3.3625	3.9827	4.6628
9	3.12833	3.9708	4.6880	*5.4823 (4535/4536%) 5.4819	2.87216	3.66305	4.3330	5.0940

* The exact significance probability is given within braces in percentage form. To obtain the α marked at the top of the column, randomization is required between the cut off points given above and below the percentage.

In order to investigate the accuracy of the large sample approximation, we evaluate both the normal approximation and the Edgeworth expansion containing the first non zero correction term. The latter expansion assumes the form (see Cramér [4] pp. 88)

$$(2.3) \qquad P\left(\frac{T_n}{\sigma_o(T_n)} < x\right) = \Phi(x) + \frac{\Gamma_{4n}}{n\Gamma_{2n}^2(4!)}\Phi^{(4)}(x) + O(n^{-2})$$

where Φ is the standard normal cdf, $\Phi^{(4)}$ is its fourth derivative, $\sigma_o^2(T_n)$ is the variance and $n\Gamma_{\nu n}$ is the νth cumulant of T_n under the null hypothesis. By symmetry of the null distribution the odd cumulants are zero for both $T_n^{(1)}$ and $T_n^{(2)}$. For $T_n^{(1)}$, we have

$$(2.4) \qquad \begin{aligned} \Gamma_{4n} &= n^{-1}\sum_{j=1}^{n} b_{nj}^4 [j^{-1}\sum_{k=1}^{j} c_{jk}^4 - 3j^{-2}(\sum_{k=1}^{j} c_{jk}^2)^2], \\ \sigma_o^2(T_n^{(1)}) &= n\Gamma_{2n} = \sum_{j=1}^{n} b_{nj}^2 j^{-1}\sum_{k=1}^{j} c_{jk}^2 . \end{aligned}$$

The corresponding expressions for $T_n^{(2)}$ are obtained by setting the b_{nj}'s equal to one. The results appear in Table 2 and they show that both approximations are very good even for small samples.

Table 2. Normal and Edgeworth approximations to the significance probabilities of $T_n^{(1)}$ and $T_n^{(2)}$

n	$T_n^{(1)}$				$T_n^{(2)}$			
	critical value	Significance probability			critical value	Significance probability		
		True	Normal	Edgeworth		True	Normal	Edgeworth
6	2.9985	.05	.0482	.0488	2.6973	.05	.0502	.0507
6	3.5295	.025	.0252	.0237	3.1775	.025	.0265	.0255
7	3.3064	.05	.0509	.0516	3.0500	.05	.0495	.0499
7	3.9001	.025	.0268	.0258	3.6016	.025	.0257	.0249
8	3.6579	.05	.0499	.0504	3.3625	.05	.0498	.0501
8	4.2890	.025	.0268	.0260	3.9827	.025	.0255	.0249
9	3.9708	.05	.0498	.0502	3.66305	.05	.0497	.0499
9	4.6880	.025	.0259	.0252	4.3330	.025	.0256	.0251

3. Some dependence alternatives and empirical power

The following five tests of independence are included in the present study for the evaluation of their relative performance in small samples under specific types of dependence.

(i) The correlation coefficient r.

(ii) Spearman's rank correlation $r_s = 12(n^3-n)^{-1}\sum_{j=1}^{n}(j-\frac{n+1}{2})R_j$, where $R_1,\dots,R_n$ are the ranks of $Y_{[1]},\dots,Y_{[n]}$.

(iii) Kendall's statistic $t = [\binom{n}{2}]^{-1}\sum_{j=1}^{n}[\ell(j)-(j+1)/2]$.

(iv) The asymptotically locally most powerful normal score layer rank test $T^{(1)}$ defined by (1.1) and (1.2).

(v) A simplified normal score layer rank test $T^{(2)}$ defined in (1.3).

Three sample sizes $n=5,7,9$ and two significance levels $\alpha=.05$ and $.025$ are considered for each test. The one-sided cut-off points of $T^{(1)}$ and $T^{(2)}$ are obtained from Table 1, those of r_s and Kendall's t are read from [8] and for the cut-off points of r, we use the fact that the null distribution of $(n-2)^{1/2}r/(1-r^2)^{1/2}$ under normality is student's t with $(n-2)$ degrees of freedom.

The following three essentially different types of positive dependence are included in the study.

(a) X and Y have the bivariate normal distribution with positive correlation coefficient ρ. The tests r and $T^{(1)}$ are optimal for this family in two different senses. It is interesting to investigate the comparison of the two as well as the manner in

which they compare with the other three tests which are not optimal for this family. All the five tests are location and scale invariant. Without loss of generality we take the means to be zero, variances arbitrary and $\rho=.1$, .3, .5, .7 and .9 to cover evenly the whole range of positive dependence.

(b) A model of positive dependence which often arises in factor analysis:

$$X = V+\Theta Z \qquad (3.1)$$
$$Y = W+\Theta Z$$

where V, W and Z are independent random variables and Θ is the parameter of positive dependence. The null hypothesis is equivalent to $H_o:\Theta=0$. This model was considered by Bhuchongkul [2] and also by Hájek and Šidák [6] in the derivation of locally optimal rank tests for independence. The particular case of V, W and Z being independent uniform random variables on (0,1) is treated here. The correlation coefficient between X and Y is $\rho=\Theta^2/(1+\Theta^2) \geq 0$. The values of Θ are selected to yield alternatives which correspond to the values of ρ in case (a).

(c) The bivariate exponential distribution introduced by Marshall and Olkin [11]. It is essentially a three parameter family with the cdf determined from

$$P(X>x,Y>y) = \exp[-\lambda_1 x-\lambda_2 y-\lambda_{12}\max(x,y)], \qquad (3.2)$$
$$x,y\geq 0,\ \lambda_1>0,\ \lambda_2>0,\ \lambda_{12}\geq 0.$$

The X and Y marginals are univariate exponential with means $(\lambda_1+\lambda_{12})^{-1}$ and $(\lambda_2+\lambda_{12})^{-1}$ respectively and the correlation coefficient is given by $\rho=\lambda_{12}(\lambda_1+\lambda_2+\lambda_{12})^{-1}$. Marshall and Olkin [11] showed that the distribution (3.2) arises very naturally in certain life testing situations. It has an interesting feature in that although the marginals have exponential distributions, the bivariate distribution has a singular part in addition to an absolutely continuous part. To reduce the number of parameters, we consider the case of identical marginals $\lambda_1=\lambda_2=\lambda_o$, say. Since all the tests are scale invariant, their powers would depend only on $\Theta=\lambda_{12}/\lambda_o$ and hence, without loss of generality, we take $\lambda_o=1$ and $\lambda_{12}=\Theta$. Then we have $\rho=\Theta/(\Theta+2)$. The alternatives Θ are again chosen so as to have the same values of ρ as mentioned in (a).

10,000 samples are generated in each case for the evaluation of empirical power. To reduce the time for generating the samples and also to provide a possibly better basis for comparison, nine pairs of observations (X_i,Y_i) are generated each time. The first five are used for the case n=5, the first seven for the case n=7 and all nine for the case n=9. As a preliminary to generating the (X,Y) observations, independent uniform (0,1) random numbers are first generated following the scheme discussed in Moshman [12]. Appropriate transformations are then used to convert to the distributions specified in (a), (b) and (c). Specifically, let Z_1, Z_2 and Z_3 be three independent uniform (0,1) random variables and set

$$(3.2)\quad X^{(a)} = (-2\ln Z_1)^{1/2} \sin(2\pi Z_2)+\rho(1-\rho^2)^{-1/2}(-2\ln Z_1)^{1/2}\cos(2\pi Z_2)$$

$$Y^{(a)} = (-2\ln Z_1)^{1/2} \cos(2\pi Z_2),$$

$$(3.3)\quad X^{(b)} = Z_1+[\rho/(1-\rho)]^{1/2}Z_3$$

$$Y^{(b)} = Z_2+[\rho/(1-\rho)]^{1/2}Z_3$$

and

$$(3.4)\quad X^{(c)} = \min[-\ln(1-Z_1),\ -(1-\rho)/(2\rho)\ln(1-Z_3)]$$

$$Y^{(c)} = \min[-\ln(1-Z_2),\ -(1-\rho)/(2\rho)\ln(1-Z_3].$$

It is easy to verify that $(X^{(a)}, Y^{(a)})$ is a bivariate normal, $(X^{(b)}, Y^{(b)})$ conforms to the model (b), $(X^{(c)}, Y^{(c)})$ has a bivariate exponential distribution and the correlation in each case is ρ. The transformation (3.2) is due to Box and Muller [3] and (3.4) follows from the results in [11].

Independent sets of uniform random numbers are used for each of the three different models. However for each model, the same set of uniform random numbers are used to generate each set of the (X,Y) observations corresponding to an alternative ρ. The observations in any set are obtained by merely changing the value of ρ in the transformations introduced above. The number of rejections of H_o out of a total of 10,000 samples, are presented in Table 3 for model (a), Table 4 for model (b) and Table 5 for model (c). Hence the entries in these tables need only be divided by 10,000 to yield the estimated power. Finally, the number of rejections are also given for $\rho=0$ which is equivalent to the hypothesis of independence in each model.

This provides a check on the accuracy of the empirical type I error.

Table 3. Number of rejections out of 10,000 trials under dependence alternatives in model (a).

Test	n	$\alpha = .05$						$\alpha = ,025$					
		ρ						ρ					
		0	.1	.3	.5	.7	.9	0	.1	.3	.5	.7	.9
r	5	454	652	1225	2293	4202	8078	244	313	633	1306	2720	6735
	7	486	727	1634	3304	6181	9494	224	373	926	2108	4678	8994
	9	512	848	2045	4198	7430	9874	233	421	1204	2949	6203	9713
$T^{(1)}$	5	471	636	1120	1916	3320	5769	226	296	554	998	1946	3801
	7	515	733	1411	2745	5033	8563	241	370	811	1705	3558	7421
	9	506	780	1774	3535	6414	9557	264	408	1010	2379	4978	9031
$T^{(2)}$	5	475	641	1105	1877	3259	6260	248	328	579	1056	1906	4222
	7	522	729	1388	2628	4805	8323	252	385	794	1633	3405	7154
	9	516	766	1683	3331	6046	9355	266	413	958	2165	4662	8772
r_s	5	476	629	1098	1874	3308	6342	237	312	567	1027	1926	4214
	7	495	713	1437	2758	5025	8564	235	357	778	1662	3511	7394
	9	499	775	1781	3575	6449	9565	271	404	1032	2404	4973	9071
t	5	475	630	1096	1874	3274	6301	237	312	567	1027	1926	4214
	7	508	716	1415	2712	4940	8476	240	369	785	1652	3465	7306
	9	499	772	1745	3502	6353	9521	263	409	1019	2361	4944	9034

Table 4. Number of rejections out of 10,000 trials under dependence models in model (b).

Test	n	$\alpha = .05$						$\alpha = .025$					
		ρ						ρ					
		0	.1	.3	.5	.7	.9	0	.1	.3	.5	.7	.9
r	5	498	663	1127	2053	4080	8482	256	332	601	1126	2530	7029
	7	501	767	1512	3069	6218	9721	261	390	853	1906	4512	9334
	9	504	793	1785	3938	7733	9963	261	432	1061	2609	6284	9891
$T^{(1)}$	5	517	641	1070	1773	3155	6219	246	318	569	977	1857	4125
	7	521	746	1391	2574	4898	8644	274	387	780	1536	3339	7432
	9	530	790	1623	3363	6461	9668	255	441	949	2129	4811	9182
$T^{(2)}$	5	507	625	1062	1783	3153	6223	269	322	542	979	1769	4120
	7	542	713	1335	2451	4626	8397	257	371	770	1463	3196	7202
	9	534	765	1588	3108	5993	9471	256	403	902	2000	4429	8876
r_s	5	509	632	1057	1792	3203	6335	258	320	556	978	1813	4123
	7	515	722	1350	2626	5114	8841	264	365	767	1570	3522	7690
	9	500	771	1616	3504	6796	9732	257	421	939	2239	5188	9361
t	5	512	628	1058	1773	3153	6257	258	320	556	978	1813	4123
	7	520	720	1361	2552	4936	8662	262	364	761	1515	3359	7469
	9	508	770	1608	3333	6507	9680	251	419	930	2175	4951	9271

Table 5. Number of rejections out ot 10,000 trials under dependence alternatives in model (c).

Test	n	α = .05						α = .025					
		ρ						ρ					
		0	.1	.3	.5	.7	.9	0	.1	.3	.5	.7	.9
r	5	713	1084	2216	3917	6060	8664	406	693	1532	3092	5308	8318
	7	731	1224	2607	4600	6832	9039	434	771	1917	3761	6125	8730
	9	731	1229	2879	5098	7447	9342	450	824	2159	4260	6756	9064
$T^{(1)}$	5	517	867	2001	3755	6029	8654	246	428	1114	2473	4559	7842
	7	521	997	2638	5012	7472	9403	274	554	1759	3904	6616	9132
	9	530	1099	3224	6011	8509	9838	255	622	2270	4918	7806	9683
$T^{(2)}$	5	507	853	2014	3810	6109	8702	269	462	1306	2799	5088	8177
	7	542	1034	2790	5237	7697	9499	257	573	1919	4159	6842	9228
	9	534	1150	3385	6243	8582	9827	256	670	2460	5195	7969	9716
r_s	5	509	829	1937	3640	5862	8545	258	445	1210	2636	4824	8010
	7	515	989	2457	4483	6788	8974	264	532	1641	3586	6201	8873
	9	500	1063	2934	5431	7830	9526	257	625	2039	4335	6899	9175
t	5	512	840	1964	3711	5971	8614	258	445	1210	2636	4824	8010
	7	520	996	2653	5028	7526	9471	262	548	1769	3892	6564	9080
	9	508	1092	3229	6095	8581	9863	251	628	2315	5054	7950	9765

4. Remarks and Conclusions

It was felt that the generated data were quite reliable since the histograms of the univariate marginal observations were well within the limits of the goodness of fit tests in each case. Moreover, to check the extent of internal variability of the results, the program was run in ten groups of 1,000 samples each. The variation of the results of empirical power from group to group was found to be quite small. This observation enhances the reliability of the pooled empirical power presented in the tables of Section 3. Further, the results for r were compared with selected values of the exact power under normal alternatives (see David [5]) and were found to agree within $\pm$ .01.

Among the five tests included in the study, all but the correlation coefficient test r are distribution-free. As mentioned earlier, the cut-off points for r were taken from student's t distribution which holds only under the bivariate normal model. The empirical significance probabilities for r in Table 3 are as close to the nominal levels as they should be. Curiously enough, they also tend to agree very closely in Table 4. The underlying model in this case is the uniform distribution of (X,Y) on the unit square. Thus, the significance levels for the parametric test r seem to be robust with respect to the uniform distribution. However, this property does not hold for the bivariate exponential model. This is evident as one examines the empirical levels of r in Table 5.

Finally, we make a comparative study of the empirical powers of the five tests under consideration. Regarding the bivariate normal model (a), there is clear indication that r has substantially higher power than all the other tests (see Table 3). r is even noticeably better than the optimal layer rank test $T^{(1)}$, although asymptotically $T^{(1)}$ has Pitman efficiency 1 relative to r. The r test also seems to be superior to all the others for the model (b). However, for the bivariate exponential alternatives, not only are its significance levels seriously distorted, but its power tends to lag behind those of all other tests. Thus among the five tests, the r test is the best for the models (a) and (b), but it is the poorest for model (c). The simplified normal score layer rank test $T^{(2)}$, on the other hand, seems to have the highest power among all five for the model (c), although it has the lowest powers for (a) and (b). The difference among the performances of $T^{(1)}$, r_s and t are somewhat less pronounced. Asymptotically, r_s and t have relative efficiency 1 for all parent bivariate distributions and their efficiency relative to $T^{(1)}$ is $(3/\pi)^2 = .912$ for the bivariate normal model (a). In Table 3, their empirical powers are found to be quite close to one another for all the alternatives ρ. For the model (b), the Spearman test r_s seems to have more power than both $T^{(1)}$ and t. The difference, though moderate, is noticeable for all n, all alternatives ρ and for both the values of α. On the other hand, r_s has slightly lower power than $T^{(1)}$ and t for the model (c).

References

[1] Barndorff-Nielsen, O. and Sobel, M. "On the distribution of the number of admissible points in a vector random sample", Theory of probability and its applications, 11 (1966) No. 2, 249-269.

[2] Bhuchongkul, S. "A class of nonparametric tests for independence in bivariate populations", Ann. Math. Statist., 35 (1964), 138-149.

[3] Box, G.E.P. and Muller, M. "A note on the generation of random normal deviates", Ann. Math. Statist., 29 (1958), 610-611.

[4] Cramér, H. Random variables and probability distributions. (2nd. ed.) (1962) Cambridge Univ. Press, London.

[5] David, F.N. Tables of the ordinates and probability integral of the distribution of the correlation coefficient in small samples. Biometrika (1938), London.

[6] Hájek, J. and Šidák, Z. Theory of rank tests. (1967), Academic Press, New York.

[7] Kaarsemaker, L. and Van Wijngaarden, A. "Tables for use in rank correlation", Statistica Neerlandica, 7 (1953), 41-54.

[8] Kendall, M.G. Rank correlation methods. (3rd ed.) (1962), Griffin & Co., London.

[9] Konijn, H.S. "On the power of certain tests for independence in bivariate populations", Ann. Math. Statist., 27 (1956), 300-323.

[10] Lehmann, E.L. "Some concepts of dependence", Ann. Math. Statist., 37 (1966), 1137-1153.

[11] Marshall, A.W. and Olkin, I. "A multivariate exponential distribution", Jour. Amer. Stat. Assoc., 62 (1967), 30-44.

[12] Moshman, J. "The generation of pseudo-random numbers on a decimal calculator", Jour. of the Assoc. for Computing Machinery, 1 (1954), 88-91.

[13] Woodworth, G.G. "On the asymptotic theory of tests of independence based on bivariate layer ranks", Tech. Report No. 75, (1966), Univ. of Minnesota.